DOCUMENT RESUME

ED 063 262 SP 005 726

AUTHOR Giacquinta, Joseph B.; And Others
TITLE The Study of Paraprofessionals in Urban Schools: Gaining Research Mileage Out of Educational Evaluations.
SPONS AGENCY New York Univ., N.Y. School of Education.
PUB DATE Apr 72
NOTE 39p.; Paper presented at the annual meeting of the American Educational Research Association (Chicago, Ill., April 1972)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Administrator Attitudes; *Educational Innovation; *Organizational Change; *Paraprofessional School Personnel; *Teacher Attitudes; *Urban Schools

ABSTRACT

Paraprofessional programs in the schools of a metropolitan district were studied to test a theory about factors related to successful implementation of organizational innovation. Indices of variables were constructed from items on two questionnaires administered to a stratified random sample of paraprofessionals, teachers, and administrators. The analysis showed how paraprofessional clarity, willingness, ability, resources, and school compatibility were predictive of differences in the implementation of the programs. The study demonstrates how data gathered during educational evaluations can be used to explore general social science theories that are of value to education. (Author)

FILMED FROM BEST AVAILABLE COPY

ED 063262

The Study of Paraprofessionals in Urban Schools: Gaining Research Mileage Out of Educational Evaluations*

Joseph B. Giacquinta, Herbert I. London,
and Irene S. Shigaki

New York University[1]

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Very few planned domestic efforts in this country have surpassed those made by metropolitan school systems during the last decade. Massive amounts of time, money, and energy have been spent on the creation and introduction of innovative programs into inner-city schools to offset the poor academic performance of children in them.[2] These efforts have been accompanied by the literally thousands of studies, massive in their own right, designed to measure the effects of these innovative programs and by a growing research literature devoted to articles and books on how to assess the effectiveness of educational innovations.[3]

While the literature and many of the studies contained in it discuss at length the problems and procedures of evaluations, surprisingly little attention is given to how data collected during evaluations may be used to increase the understanding of the educative process and to shed light on social and psychological theories related to education. The plain fact is that, since this has not been of central concern to evaluators, theory related to education has benefitted marginally from these thousands of studies.[4]

---

*A paper to be presented at the 1972 annual meeting of the American Educational Research Association in Chicago, April 1972.

SP005726

One reason for this sad state of affairs is that the conditions under which evaluation data are collected are often inadequate for research purposes and this renders them useless as the basis for developing or confirming theories and hypotheses. A second reason is that, even when collected under acceptable conditions, data from one study cannot be related in any systematic fashion to data gathered in another, so that no benefit can be derived by accumulating results. A third reason is that investigators fail to treat evaluation studies as opportunities to also collect evidence in actual field settings to explore meaningful theoretical problems. It is with this objective in mind that the research reported in this paper was undertaken.

The Social Science Problem to be Explored

During an evaluation study, data were collected to test the usefulness of a theoretical model, which includes primary conditions necessary for the implementation of proposed organizational changes in schools and other settings such as businesses, factories, and hospitals. Implementation may be defined as a process leading to the changed role performance of organizational members after the introduction of an innovation.[5] Evaluation studies, such as this one, provide excellent opportunities to explore this theoretical problem in school settings.

During the 1970-71 academic year, an assessment of educational programs funded under Title I of ESEA in a New York City School District was conducted. The district was fairly representative of inner-city areas: housing varied from decrepit to acceptable; most families were lower class and non-white; and many were on public assistance.

The primary objective of the evaluation study was to measure the effects of these programs (10 in all) on the schools and the students within them. Five of the programs in the District's 21 schools were designed specifically to offset the poor academic performance of students by providing greater individual instruction through the use of paraprofessional assistance. These five programs varied in size: the Preprimary Program involved nearly 45 paraprofessionals in early childhood and kindergarten classrooms in 10 schools; the Early Primary Program had some 160 paraprofessionals located in first and second grade classes in 11 schools; the Late Primary Program involved 35 paraprofessionals in third grade classrooms in 11 schools; the Intermediate Program contained 75 paraprofessionals in fourth, fifth, and sixth grade classes in 8 schools; and 75 paraprofessionals were assigned to classrooms in the four junior high schools.[6]

The paraprofessionals, themselves, were nearly all non-white women, 30 to 50 years of age. All had at least high school diplomas, supplemented by District-sponsored training. Many were enrolled in community college programs, with some working toward their bachelor's degree.

The five programs were treated as organizational innovations because they involved the introduction of a new position in schools at the classroom level, which resulted in a new set of role relationships. Although similar in many respects, each program was treated as a discrete innovation because the role of the paraprofessional differed by grade level. For example, the Preprimary paraprofessionals were expected to spend much of their time on activities that would promote child socialization and much less time on developing reading skills; the reverse was true for the paraprofessionals in the Junior High Program.

Pertinent Social Science Literature

The study of planned change in the past has emphasized several areas. In their writings, Carlson[7], Lionberger[8], and Rogers[9] underscore the interest many investigators have in explaining the rates at which simple, technological innovations are diffused and adopted among individuals -- e.g., hybrid seed among farmers, hand tools among villagers, tranquilizers among doctors, and audio-visual aids among school superintendents (representing school systems). Some of the major questions examined by researchers have been: What are the stages that individuals go through in the adoption process? How do characteristics of the innovations influence their rates of diffusion? What personal and social attributes distinguish early from late adopters?

The work of Bennis, Benne, and Chin[10] and Lippitt, Westley, and Watson[11] reflect an area of interest of still other investigators, one more clearly associated with schools as organizations: the initiation of organizational changes -- the process whereby innovations are introduced and adopted in organizations[12]. Here a major concern has been to determine to what extent the use of tactics such as hiring change agents, permitting subordinates to participate in decision-making, or both, leads to successful initiation of changes.

A recent appraisal of the planned change literature by Gross, Giacquinta, and Bernstein[13], however, reveals that the problem of implementing organizational innovations, as compared to the problem of initiation, has received very little systematic empirical study. This study notes that much of the discussion of implementation usually relates to the potential resistance of members toward changes which are introduced, and to the conditions necessary for overcoming this resistance. It calls

into question a number of the assumptions underlying this reasoning about resistance and points to the critical need for studying implementation as a process separate from initiation.[14]

The theoretical model describing the factors that are at the heart of implementation and emerging from this recent work, serves as the basis for this present investigation in schools. Gross and his associates propose that the implementation of organizational innovations having once been initiated requires the presence of five conditions, each of which is necessary but alone insufficient. These antecedent conditions of implementation are that: 1) organizational members who must carry out an innovation clearly understand the new expectations; 2) they have the abilities necessary for the new role performance; 3) they are willing to make the required efforts to behave in accord with the innovation; 4) the materials and resources required by the innovation are available; 5) the organization is compatible and supportive of the innovation[15].

According to this model, the degree to which innovations are implemented is a function of the extent to which these conditions in combination are present. Thus, programs or schools that rank higher on these conditions should rank higher on implementation than those that rank lower on these conditions. Moreover, programs or schools (containing programs) where the five conditions are highly present should demonstrate high levels of implementation. The evaluation study, from which this investigation emanates, permitted the collection of data that could be used to test the usefulness of this model in explaining variations in level of implementation among five innovative paraprofessional programs and among several schools within two of these five programs.

Procedures

Sample

Since the number of paraprofessionals in each program varied, a proportionate, stratified random sample was drawn. From the total of approximately 390 paraprofessionals working in these programs, 127 were selected. After the sample of paraprofessionals was chosen, the teacher with whom each worked was identified. The original total sample for the study, therefore, was 254 individuals, half of whom were paraprofessionals and the other half teachers. A number of administrators in each program were also selected and used primarily for purposes of research beyond the scope of this study.

Data Collection

Two instruments were employed in the collection of data. The purpose of the Role Definition Questionnaire (RDQ) was two-fold: First, it provided evidence about the role definition of the paraprofessional to be used in assessing implementation; second, it provided information for one of the independent variables, clarity of role. The instrument is comprised of 56 items that reflect activities often related to the role of the paraprofessional.

Five trained research assistants administered the RDQ during March, 1971 to the 127 paraprofessionals in our sample. Employed for the on-going evaluation of the paraprofessional programs, the assistants, in most cases, had previous contact with the subjects. The questionnaire, accompanied by thorough instructions, was personally delivered to each subject. Arrangements were made for the questionnaires to be collected by the research assistants. In the few cases where questionnaires were not completed, subjects were asked to return them by mail. The personal

contact by the research assistants contributed to the high rate of returns, which ranged from 80 to 100 per cent (see Table 1).

The second instrument constructed for the present study was the Implementation Questionnaire (IMQ). The purpose of the instrument was to gather demographic information, evidence on the degree of implementation, and data on four of the independent variables: ability, willingness, availability of resources and materials, and organizational compatibility. The instrument was administered in May, 1971 to the sample of paraprofessionals and teachers as well as to some administrators. As mentioned earlier, the expected sample size was 254. However, some attrition of paraprofessionals occurred between the administration of the RDQ and the IMQ due primarily to reassignment. Since the paraprofessionals and teachers were selected together, the attrition rate in the original sample was doubled. The same research assistants administered the IMQ in the manner described above. Returns for each program ranged from 74 to 100 per cent (see Table 1).

Consistency of Respondents: To obtain a measure of the respondents' consistency, ten of the original 56 items on the RDQ were randomly selected, reworded, and included at the end of the questionnaire. Consistency was considered to reflect the seriousness with which the paraprofessionals viewed their responses to the RDQ. Correlations between responses to the original items and the paraphrased versions ranged from .10 to .53 and are listed in Table 2. All but one of the ten correlations were significant, although they were not as high as had been anticipated.

Several reasons may account for the rather low correlations, but these reasons do not reduce the confidence in the accuracy of paraprofessional responses to the questionnaire. First, to guard against a response set, the direction of the scale for the ten reworded items was reversed. In addition, the respondents were asked to record rather than circle their choices. In spite of careful instructions, many paraprofessionals could have simply continued using the older response scale thereby lowering the correlations. Second, although the consistency items were carefully paraphrased to maintain the content of the original ones, the rewording of the items could have introduced dimensions not present in the originals, thereby influencing the responses to them and lowering the correlations. Third, since the consistency items were attached to the end of a 56-item questionnaire, fatigue may have taken its toll resulting in less careful selections.

Inspection of the questionnaires themselves and the paraprofessionals' comments to the assistants who gathered them revealed that a great deal of concern and effort were given to the answers. Nonetheless, it is still possible that the low correlations reflect inconsistencies of respondents due to several reasons: the level of reading comprehension by the paraprofessionals may have been inadequate; the ratings may have required too fine a discrimination by the paraprofessionals; and, the effort of the paraprofessionals may have been minimal resulting in haphazard responses.

Measurement of Variables

The utility of this theoretical model was tested with three sets of analysis. The first set involved comparisons among the five programs (some cutting across as many as 11 schools). The second set required comparisons of the four schools into which the Junior High Program was introduced. The third, also a comparative school analysis, involved six schools having the Early Primary Program. The two school analyses were possible because two programs contained large concentrations of paraprofessionals in many schools. Therefore, when the original stratified sample of paraprofessionals was drawn for the program comparisons, enough paraprofessionals (and their teachers) were included in these schools to develop an adequate summary school score for each of the five independent variables and the dependent variable. The procedures that follow were used to obtain the summary scores.

Degree of Implementation: Fifteen items about the paraprofessionals' role performance were included on the IMQ in order to assess the degree of actual implementation. The items emerged from our analysis of the paraprofessionals' responses on the RDQ, program documents discussing the nature of the innovation, and our professional judgments. Ten of the items used for the Early Primary Program, for example, described behavior that paraprofessionals in this program would be exhibiting if they were properly carrying out their role and five activities that they would not exhibit if they were properly implementing the role. The "positive" dimensions were:

1) work with children in classrooms on reading skills.

2) work with children in classrooms on reading comprehension.

3) work with children in classrooms on math skills.

4) work with children who need remedial work in reading.

5) work primarily with small groups of children on school subjects.

7) plan classroom activities with the teacher(s).

9) perform clerical duties within the classroom such as taking attendance, correcting papers, preparing and duplicating materials.

12) confer with teachers about children.

14) attend training sessions for paraprofessionals.

15) keep abreast of materials on education that will be helpful on the job.

The five inappropriate activities were:

6) work primarily on non-instructional matters with children in the classroom.

8) act as a substitute when a teacher is absent from school.

10) help to prepare snacks and lunches for children.

11) help in the health clinic to care for those children in need of medical attention.

13) visit children and their parents at home regarding instructional and/or non-instructional matters.

The teachers were asked, as observers, to estimate how frequently their paraprofessionals were engaging in each of these 15 activities using the following scale: 1=never, 2=seldom, 3=occasionally, 4= usually, 5=always, N=cannot judge. The paraprofessionals were

asked to assess other paraprofessionals, with whom they had contact, on these 15 activities. Weighting the items equally, the responses of each teacher and paraprofessional were summed (after reversing the scoring of the negative items) and divided by 15. The resulting mean scores represented the subjects' overall assessments of the degree to which paraprofessionals were carrying out the role. Subjects' mean scores were then averaged either across programs or schools (depending upon the analysis). These mean scores, in turn, represented summary indices of implementation for programs or schools.

Employing different sets of items, this procedure was followed to obtain summary scores for programs and schools on four of the independent variables: the capability of paraprofessionals, their willingness, the availability of necessary resources, and organizational support. Since these variables are specific for the innovation being studied, establishing the validity of the instrument is especially difficult. One way of providing evidence on the validity of the measure, although not without apparent limitations, would be to see if the results of distinct groups using the instrument are similar.

For each program the mean scores of teachers, paraprofessionals, and some administrators (who were given the IMQ for other purposes) on the five conditions were separated and subjected to analyses of variance to determine whether assessments of these conditions differed among the three groups. The results are summarized in Table 3. Only two of the 25 ANOVA's indicated significant mean differences between teachers and paraprofessionals: the ANOVA on Early Primary resources and the ANOVA on Intermediate willingness. The similarity of results obtained from the three groups measured independently using the same instrument,

adds confidence to the validity of our measures. Once these results were obtained, it became apparent that the inclusion of the data from the administrators introduced an unnecessary encumbrance on the research analysis. Therefore, they were omitted at this point, and the scores of teachers and paraprofessionals were pooled. The scores of the teachers and paraprofessionals were pooled in order to arrive at the final summary scores to be used in the program and school analyses. The summary mean scores for the five programs are reported in Table 4, for six of the schools in the Early Primary Program in Table 5, and for the four schools in the Junior High Program in Table 6.

The summary program scores on implementation as evidenced in Table 4 ranged from 4.04 to 3.65. On a scale from 1 (low) to 5 (high), this can be interpreted to mean that overall the five programs were being implemented at a fairly high level and that the variation between them was not great. The summary school scores for the six Early Primary schools (Table 5) ranged from 4.19 to 3.60 and in the Junior High schools (Table 6) from 4.17 to 3.52.

Paraprofessional Ability to Carry Out the Role: To obtain a measure of the degree of paraprofessional ability in each program, only the "positive" items used in measuring implementation were employed. For example, in the Early Primary Program respondents were asked to judge how capable were the paraprofessionals to perform the following activities:

1) working with children in classrooms on reading skills.

2) working with children in classrooms on reading comprehension.

3) working with children in classrooms on math skills.

4) working with children who need remedial work in reading.

5) working primarily with small groups of children on school subjects.

6) planning classroom activities with the teacher(s).

7) performing clerical duties within the classroom such as taking attendance, correcting papers, preparing and duplicating materials.

8) conferring with teachers about children.

9) attending training sessions for paraprofessionals.

10) keeping abreast of materials on education that will be helpful on the job.

They were asked to use the following code in responding: 1=completely incapable, 2=somewhat incapable of, 3=somewhat capable of, 4=very capable of, 5=completely capable of, and N=No basis for judging. The summary scores for programs and schools reported in Tables 8, 9, 10 were computed in the manner employed in measuring implementation. Variation in ability by program ranged from 4.54 to 3.47. Variation in ability by school in the Early Primary Program ranged from 4.47 to 2.91 and in the Junior High Program from 3.76 to 2.51. On a five-point scale from 1 (low) to 5 (high) the variation between schools within the two programs was much greater than across projects, even though ability was still generally high across programs.

Willingness of Paraprofessionals: To assess the degree to which paraprofessionals in each of the programs were willing to implement their role, the positive items, as used in the ability section, were presented with the following code: 1=never willing to, 2=seldom willing to, 3=occasionally willing to, 4=usually willing to, 5=always willing to,

and N=no basis for judging. Variation in willingness across programs ranged from 4.70 to 4.02 (Table 8). In the Early Primary schools the variation went from 4.89 to 4.12 (Table 9) and in the Junior High schools from 4.54 to 3.96 (Table 10). Willingness also appeared to be generally high.

Availability of Necessary Resources and Materials: To measure the adequacy of resources and materials respondents were asked to judge the adequacy of a) materials, b) physical facilities, c) financial resources, and d) staff, on a five-point scale from 1=inadequate, 2=somewhat inadequate, 3=somewhat adequate, 4=adequate, 5=completely adequate. Each subject's mean score based on these four items was calculated and then similar calculations were made for programs and schools. Compared to ability and willingness, the availability of resources across programs and schools was lower. Mean program scores ranged from 3.30 to 2.47 (Table 8), in the Early Primary program by school from 3.40 to 1.43 (Table 9), and in the Junior High schools from 2.72 to 1.96 (Table 10).

Organizational Compatibility: Seven sources of organizational incompatibility were identified for this study and put into question form. Subjects were asked to judge to what extent standard procedures or policies in schools were interfering with the installation of the new paraprofessional role. Given the following code -- 1=no and 2=yes -- they were asked to answer the following questions for the paraprofessionals:

1) Do expectations or policies of their school administrators interfere with their carrying out the role?
2) Do expectations or policies of their teachers interfere with their carrying out the role?

3) Do demands made of them by students get in the way?

4) Do demands of parents interfere?

5) Do demands of the community interfere?

6) Are there any district policies affecting schools that act as obstacles to their carrying out the role?

7) Are there expectations of personnel in the program which interfere with their effective performance?

A summary score for each subject based on these items was computed. Scores could range from 1.00 (compatibility) to 2.00 (incompatibility). Program scores ranged from 1.00 to 1.03 (Table 8), Early Primary school scores varied from 1.00 to 1.13 (Table 9), and Junior High school scores from 1.00 to 1.14 (Table 10). With high compatibility as 1.00, there was almost no variation either by program or by schools within programs. High compatibility appeared to be present.

Independence of Staff Judgments on the IMC: The results may be contaminated by the fact that subjects were automatically judging everything as "high" or as "low". To determine the independence of the respondents' judgments on the five variables discussed above, their mean scores were correlated (Table 7). The intercorrelations were computed for each program. Although many of the correlations between implementation and the four independent variables were significant, in only two (both in the same program) is more than 25 per cent of the variance in implementation explained. The strongest consistent correlation among the independent variables was that between ratings on ability and on willingness, e.g., in the Preprimary Program the correlation was .92, the Early Primary .83, Late Primary .89, Intermediate .31 and Junior High School .68. The

weakest consistent correlations were between ability and compatibility and between resources and compatibility. The findings suggest that the respondents' ratings on the variables were, in the majority of instances, substantially independent of each other.

Clarity of the Paraprofessional Role: The degree to which the paraprofessionals were clear about the innovation and their role was determined in the following manner. It was reasoned that the degree of consensus of the paraprofessionals on the 56-item RDQ would reflect the degree of clarity prevailing in the programs and schools. Clearly, if there was little agreement found among them about how much time they should spend on these activities, this would be evidence of lack of consensus; the lower the consensus, the lower the clarity.

It was judged that there was high paraprofessional agreement on the importance of any given activity when at least 80 per cent of the paraprofessionals' responses fell within no more than two categories (e.g., 4's and 5's or 1's and 2's). The responses for each item were: 1= a person with this job should be spending none of one's time, 2= very little of one's time, 3= some of one's time, 4= a good deal of one's time, and 5= as much of one's time as possible. Among the Preprimary paraprofessionals there was agreement on only 7 per cent of the items, among the Early Primary subjects 5 per cent, among the Late Primary respondents 9 per cent, among the Intermediate 29 per cent, and among the Junior High paraprofessionals 7 per cent (Table 8). The evidence suggests that there was a great lack of consensus, and thus lack of clarity, in each program and further, that there was little variation between programs. Variation was greater within the Early Primary schools ranging from 34 per cent to 7 per cent agreement, and within the Junior High schools from agreement on 52 per cent of the items to 7 per cent.

Results

On the basis of the summary scores computed above, the programs (and schools) were rated as high, moderately high, moderately low, or low on each of the five conditions depending upon which quartile of the scale for that variable the summary score fell. Then, composites of ratings for the programs (and schools) were formed and on the basis of these composites, the programs were ranked on implementation. For example, a program with three "high" ratings, one "moderately high" rating, and one "low" rating was ranked above a program with two "high" ratings, one "moderately high" rating, one "moderately low" rating, and one "low" rating. This predicted order, made independently of the assessed rank order based on the implementation index, was computed and compared with the assessed implementation index. A Spearman's rho was calculated, and its level of significance specified. Identical procedures were employed for the two school analyses. If, as predicted on the basis of the theory, the three analyses uncovered strong rank order correlations between the predicted rankings and the assessed rankings of schools or programs on implementation, this would be interpreted as support for the theoretical model.

Table 11 summarizes the program analysis, which, according to the theory, should reveal a strong association between the predicted rank ordering of the programs on implementation and the assessed rank ordering of the programs on implementation. As a result of the composite ratings, the Preprimary Program was ranked as number one, Early Primary as second, Intermediate as third, and the Late Primary and Junior High Programs as tied for 4th and 5th places. The rank order based on the implementation index also placed Preprimary one, Early Primary two, Junior High three, and Late Primary and Intermediate as tied for 4th and 5th places. The

correlation of .778 between these two rank orders was in the expected direction and significant at the .08 level.

The rank order analysis of six of the schools in the Early Primary Program is summarized in Table 12. The association between the predicted order of the schools on implementation and their assessed order is strong and in the expected direction. A Spearman's rho was calculated to be .872 and was significant at the .02 level.

The results of the rank order analysis of the four schools having Junior High Paraprofessional Programs are summarized in Table 13. The relation between the predicted and assessed rankings is again strong and in the expected direction. A Spearman's rho calculated to be .995 was significant at the .05 level.

In sum, the three analyses revealed strong associations and acceptable levels of significance between the expected and measured rank orders despite the size of correlations required with few cases.

Discussion and Implications

The theoretical model of implementation under examination specified that the extent to which organizational innovations are implemented depends upon the relative presence of five conditions: the clarity that organizational members, who must carry out the change, have about their role, their ability to make the necessary efforts, their willingness to do so, the availability of the necessary resources and materials, and the compatibility and support of the existing organization. It was argued that this theoretical model of implementation would be supported by the analysis if the programs or schools that ranked higher on implementation

on the basis of the five independent variables would also rank higher on assessed implementation. Since strong correlations were found between predicted and assessed rank orders in each of three analyses, the hypothesis was confirmed.

It was also argued that according to the model, schools or programs high on all five factors would demonstrate high absolute levels of implementation. This, too, was confirmed by the investigation. None of the programs or schools were high on all five conditions, and there were no examples of thorough implementation as evidenced by the absence of mean scores between 4.75 and 5.00. Close inspection of Tables 11-13 reveals that the summary implementation scores range from 3.52 to 4.19 and that nearly all programs and schools had three or more ratings of high or moderately high, while receiving low ratings on role clarity. Since the typical composite ratings, according to the model, would be predictive of substantial, though not complete, implementation, the theory was also supported in this manner. Therefore, these findings lead to the conclusion that the model has utility in explaining the implementation of organizational innovations.

Implications for Future Research: This study sheds light on conditions that accounted for variations in implementation of innovations. Further replication studies in settings where similar paraprofessional programs are introduced are needed. Moreover, to test the usefullness of the model, its utility in educational settings with innovations other than paraprofessionals are necessary, as well as settings outside of education.

Although this study isolated a number of antecedent conditions for implementation, it did not account for why they prevailed to the extent that they did, e.g., the lack of role clarity among paraprofessionals and high level of their willingness to perform their role. Research along these lines is important.

Other research must be directed to refinements in measurement, especially in the areas of reliability and validity. Though some exploratory work was done on reliability, the consistency results were not as strong as anticipated by the research team. A revision of the RDQ taking into consideration some of the possible limitations listed earlier would be desirable. Reliability of the RDQ might be assessed through correlation of responses between two randomly selected groups or a test-retest design with the same subjects. Similar reliability indices should be explored for the IMQ. In this case two randomly selected groups may be the preferred method due to the length of the questionnaire and the demands on the respondents' time. Though exploratory work on the validity of the measures in the IMQ was done, much more work is required.

Although this model yielded high predictability when compared to the implementation index, the variability of both the independent variables and implementation was restricted. Two plausible explanations are that the conditions and programs were being implemented to a similar degree or that the methodology may have contributed to the limited variance. The latter deserves further exploration. In the present study items for implementation, capability and willingness were contingent, in part, upon role definitions provided by paraprofessionals. By asking

them to define their own roles, there may have been a tendency to define it on the basis of what they were actually doing, thereby inflating the degree of assessed implementation. An alternative means of determining the parameters of a given role might be to administer the RDQ to a group familiar with the goals of the program, but who are not themselves involved in the role. Two possible groups might be the trainers of the paraprofessionals and the Title I District Administration staff. Research addressed to some of these concerns is currently in progress.

Evaluations and Social Science Research: This study was confined to the analysis of one area of social science, i.e., the implementation of organizational change. It is apparent that evaluation studies can be used to explore other areas and to test other social science theories relevant to education. The illustrations that follow are indicative of some of the kinds of research that might be pursued.

Although little systematic research has been done on the relationship between staff ethnicity and racial composition and student identification and achievement, the claim continues to be made, as several books on decentralization attest, that students who can identify with the ethnicity of their teachers will perform well and develop positive self images. Similarly, the claim is often made that parental participation in school decisions affects the learning of students. Do students with the same ethnicity as their teachers develop positive self images? Do the children whose mothers are paraprofessionals achieve more than children whose parents are not involved in school affairs? Implicit in questions such as these are psychological theories that warrant further testing.

Others, notably the United Federation of Teachers, have maintained that smaller classes will improve student achievement, especially in the area of reading. Implicit in this argument are sociological and psychological theories about the teaching and learning of reading. Evaluations of reading programs would be appropriate ways to test theories about the reading process.

Similarly, the extent to which political socialization accompanies parent participation in funded programs is an aspect of educational evaluations that often goes unnoticed. Schools are socializing agents for students, teachers and parents. And that fact is borne out most clearly in ancillary programs designed to complement the conventional school programs. When community decisions are made regarding the selection of teaching materials the components of socialization are most evident. Yet all too often evaluators assess the extent of resources, not how they are selected.

Economic matters are also subject to systematic analysis in an educational setting. If congressional debates are any index, one could easily get the impression the allocation of federal funds to local school districts is the critical determinant in predicting student achievement. But is this true? And if not, what influence do additional funds have on student performance? The issue is clearly stated as "What is the return on investment?" Another related issue is whether spending in an institution other than the school is likely to reap greater student benefits. For example, are achievement levels likely to improve more significantly if a maintenance allowance were given to the family rather than the school?

Questions, such as these, have generally not been studied by educational evaluators. Yet, large scale evaluation studies, particularly those

involving Title I, Title III, and State Urban Education funds, are vehicles for this purpose. What is needed is a clearer recognition, than has heretofore been realized, of the potential for social science research in these evaluations and its importance for education.

FOOTNOTES

[1]The research reported in this paper was sponsored by the Office of Field Research, School of Education, N.Y.U. as part of an evaluation study conducted for the Board of Education in New York City. We would like to give special thanks to Mrs. Carole Kazlow who assisted us with the computer analysis of the data.

[2]See, for example, B. J. Fallon, (Ed.), Educational Innovation in the United States, Bloomington, Ind.: Phi Delta Kappa, 1966; E. W. Gordon and D. A. Wilkerson, Compensatory Education for the Disadvantaged, (New York: College Entrance Examination Board, 1966); "Informal Education," The Center Forum, III:3 (July, 1969), pp. 2-8.

[3]See, for example, J. T. Hastings, "Curriculum Evaluation: The Why of the Outcomes," Journal of Educational Measurement, III:1 (Spring, 1966), p. 29; H. H. Hyman, C. R. Wright, and T. K. Hopkins, Applications of Methods of Evaluations: Four Studies of the Encampment for Citizenship, (Berkeley and Los Angeles, Calif.: University of California Press, 1962); C. M. Lindvall and R. C. Cox, Evaluation as a Tool in Curriculum Development, (Chicago: Rand McNally and Co., 1970); R. W. Tyler, R. M. Gagné, and M. Scriven, Perspectives of Curriculum Evaluation, (Chicago: Rand McNally and Co., 1967); E. A. Suchman, Evaluative Research: Principles and Practice, (New York: Russell Sage, 1967); J. W. Wrightstone, J. S. Coleman, D. G. Hawkridge, A. B. Chalupsky, H. S. Dyer, J. Mann, M. Mayer, E. A. Suchman, P. H. Rossi, E. Wynne, and M. Scriven, "Evaluating Educational Programs: Symposium," The Urban Review, III:4 (February, 1969) pp. 4-31.

[4]E. L. McDill, M. S. McDill, and J. T. Sprehe, Strategies for Success in Compensatory Education: An Appraisal for Evaluation Research, (Baltimore, Md.: John Hopkins Press, 1969), p. 55.

[5]For a fuller discussion of this process see, N. Gross, J. B. Giacquinta, and M. Bernstein, Implementing Organizational Innovations, (New York: Basic Books, 1971), especially pp. 1-17.

[6]A sixth paraprofessional program included in the data collection phase was dropped during the analysis because of its very small size. Three paraprofessionals comprised the total program.

[7]R. O. Carlson, Adoption of Educational Innovations, (Eugene: Oregon: Center for the Advanced Study of Educational Administration, University of Oregon, 1965).

[8]R. F. Lionberger, Adoption of New Ideas and Practices, (Ames, Iowa: Center for International Affairs, 1964).

[9]E. M. Rogers, Diffusion of Innovations, (New York: Free Press, 1962).

[10]W. G. Bennis, K. D. Benne, and R. Chin (Eds.), The Planning of Change, (New York: Holt, Rinehart, and Winston, 1961; Revised edition, 1968).

[11]R. Lippitt, J. Watson, and B. Westley, The Dynamics of Planned Change, (New York: Harcourt, Brace, and World, 1958).

[12]N. Gross, op. cit. See also M. B. Miles (Ed.) Innovation in Education, (New York: Teachers College, Columbia University, 1964).

[13]N. Gross, op. cit., pp. 19-40.

[14]Ibid.

[15]N. Gross, op. cit., pp. 195-203.

TABLE 1

Percent of Sample Returns from Five Paraprofessional Programs

Program	Role Definition Questionnaire[a]		Implementation Questionnaire[b]	
	Sample Size (N)	% Return	Sample Size[c] (N)	% Return
Preprimary	13	100% (13)	11	100% (11)
Early Primary	62	97% (60)	95	74% (70)
Late Primary	13	85% (11)	16	88% (14)
Intermediate	12	100% (12)	17	94% (16)
Junior High	27	96% (26)	42	79% (33)

[a] Paraprofessionals only

[b] Paraprofessionals and Teachers

[c] The expected sample size of the implementation questionnaire should have been double that of the role definition instrument. Because some paraprofessionals were transferred from one program to another during the year, the implementation sample size reflects this attrition.

TABLE 2

Consistency of Paraprofessionals' on the RDQ as Measured
by the Correlation of their Responses to Pairs of Parallel Items

	Role Dimension Measured by Each Pair of Items	Item Numbers on RDQ	N[a]	Correlation between items[b]
1.	Helping Students Individually with their School Work	10 & 57	97	.23*
2.	Discussing Children's Problems with Guidance Counselors	41 & 58	97	.35**
3.	Helping Children Develop Math Skills	3 & 59	105	.10 n.s.
4.	Substituting for Absent Teachers	24 & 60	99	.45**
5.	Discipling Pupils Outside Classrooms	29 & 61	101	.49**
6.	Meeting with School Administrators	40 & 62	101	.53**
7.	Participating with Planning Para-professional Activities	50 & 63	101	.26**
8.	Evaluating Student Achievement	17 & 64	93	.36**
9.	Working with Children in Non-Instructional Activities	13 & 65	99	.27**
10.	Acting as Liaison between T's and S's due to Cultural Differences	35 & 66	95	.38**

* $<.05$ (one-tailed)
** $<.01$ (one-tailed)
[a] N's vary to reflect respondents' omission.
[b] Original r's were negative since the direction of the scale for one set of items was reversed.

TABLE 3

F ratios and Significance Levels for Analyses of Variance of the Responses among Teachers, Paraprofessionals, and Administrators[a] in Each Program on Five Conditions

Program	Variable	F ratio	(d f)	Level of Significance
Preprimary	Implementation	0.4094	(2 & 42)	n.s.
	Ability	1.2179		n.s.
	Willingness	0.5308		n.s.
	Resources	1.3313		n.s.
	Compatibility	1.0497		n.s.
Early Primary	Implementation	1.4501	(2 & 70)	n.s.
	Ability	1.9790		n.s.
	Willingness	0.0224		n.s.
	Resources	4.6957		<.05
	Compatibility	0.4977		n.s.
Late Primary	Implementation	0.3938	(2 & 15)	n.s.
	Ability	0.0747		n.s.
	Willingness	0.6991		n.s.
	Resources	0.4066		n.s.
	Compatibility	0.1454		n.s.
Intermediate	Implementation	0.5978	(2 & 20)	n.s.
	Ability	0.6024		n.s.
	Willingness	4.3908		<.05
	Resources	1.7391		n.s.
	Compatibility	1.0377		n.s.
Junior High	Implementation	0.3127	(2 & 35)	n.s.
	Ability	0.3233		n.s.
	Willingness	0.0419		n.s.
	Resources	0.5909		n.s.
	Compatibility	0.0464		n.s.

[a] For the purposes of future study, the responses of various administrators in each program were included in this analysis.

TABLE 4

Summary Project Scores and Standard Deviations
For Implementation and Four Independent Variables

Program		Independent Variables				
		Ability	Willing-ness	Resources	Compat-ibility	Implemen-tation
Preprimary	$\bar{X}$	4.5364	4.7000	3.2955	1.0000	4.0381
	SD	0.5697	0.5848	0.9926	0.0000	0.3371
N = 11						
Early Primary	$\bar{X}$	4.0908	4.4321	2.6993	1.0247	3.9678
	SD	0.9130	0.8932	1.3382	0.0190	0.5730
N = 70						
Late Primary	$\bar{X}$	3.7009	4.0241	2.6875	1.0000	3.6648
	SD	1.2406	1.3712	1.5230	0.0000	1.1881
N = 14						
Intermediate	$\bar{X}$	3.6328	4.4012	2.5209	1.0060	3.6517
	SD	0.5789	0.6069	1.3009	0.0040	0.3235
N = 16						
Junior High	$\bar{X}$	3.4728	4.1406	2.4722	1.0328	3.8147
	SD	1.2930	1.2211	1.2445	0.0253	0.9007
N = 33						

TABLE 5

Summary School Scores and Standard Deviations for Implementation of the Early Primary Program and Four Independent Variables

School		Independent Variables				
		Ability	Willingness	Resources	Compatibility	Implementation
A	$\bar{X}$	4.4666	4.6534	3.4000	1.1286	4.1886
	SD	0.6485	0.3699	0.9117	0.0217	0.2472
N = 5						
B	$\bar{X}$	2.9089	4.6032	1.7583	1.0429	3.6554
	SD	0.6168	0.3699	1.2841	0.0097	1.2349
N = 10						
C	$\bar{X}$	4.6111	4.8889	3.8056	1.0000	3.9988
	SD	0.3180	0.1364	0.7265	0.0000	0.2968
N = 9						
D	$\bar{X}$	4.3800	4.4800	1.8500	1.0858	3.9190
	SD	0.0056	0.5891	0.8768	0.0128	0.3244
N = 5						
E	$\bar{X}$	2.9918	4.1205	3.1591	1.0702	3.6011
	SD	1.4384	1.4646	0.5157	0.0082	0.3732
N = 11						
F	$\bar{X}$	4.2300	4.6611	1.4286	1.0000	4.0291
	SD	0.2419	0.3131	1.3973	0.0000	0.1705
N = 7						

TABLE 6

Summary School Scores and Standard Deviations for Implementation of the Junior High Program and Four Independent Variables

School		Independent Variables				
		Ability	Willingness	Resources	Compatibility	Implementation of Junior High Program
A	$\bar{X}$	3.7636	3.9633	2.6429	1.0119	3.5373
	SD	0.6211	0.7858	1.3374	0.0101	0.3624
	N= 7					
B.	$\bar{X}$	3.5472	4.2247	2.5075	1.1346	4.1514
	SD	1.2875	1.4483	1.3025	0.0494	0.6922
	N= 11					
C	$\bar{X}$	3.9834	4.5431	2.7188	1.0000	4.1711
	SD	0.5334	0.4219	0.7372	0.0000	0.6065
	N= 8					
D	$\bar{X}$	2.5054	3.7246	1.9643	1.0000	3.5176
	SD	1.9717	1.7789	1.6100	0.0000	1.2808
	N= 7					

TABLE 7

Independence of Staff[a] Judgments about Five Conditions as Measured by the Intercorrelations of their Summary Scores

Program	Variable			
	Implementation (1)	Ability (2)	Willingness (3)	Resources (4)
Preprimary (N=15)				
2. Ability	.530*			
3. Willingness	.502*	.916*		
4. Resources	.382	.419	.439	
5. Compatibility	.072	-.375	-.298	.224
Early Primary (N=73)				
2. Ability	.475*			
3. Willingness	.495*	.833**		
4. Resources	.212	.084	-.065	
5. Compatibility	-.107	-.074	-.164	.167
Late Primary (N=20)				
2. Ability	.791**			
3. Willingness	.829**	.888**		
4. Resources	.466*	.570**	.443*	
5. Compatibility	.396*	.493*	.364	.536*
Intermediate (N=24)				
2. Ability	.349			
3. Willingness	.385	.305		
4. Resources	.152	.288	.393*	
5. Compatibility	-.085	-.074	.228	.134
Junior High (N=37)				
2. Ability	.403*			
3. Willingness	.469**	.680**		
4. Resources	.096	.252	.474**	
5. Compatibility	-.042	.101	.201	.009

* $p < .05$

** $p < .01$

[a] Includes Teachers, Administrators, and Paraprofessionals

TABLE 8

Ratings of Each Paraprofessional Program on the Five Independent Variables of Paraprofessional Clarity, Willingness and Ability, Availability of Necessary Resources, and Organizational Compatibility.

Variable	Program Ratings				
	Pre-Primary (N=11)	Early Primary (N=70)	Late Primary (N=14)	Intermediate (N=16)	Junior High (N=33)
1. Clarity of Paraprofessionals' Conception of Their Role	LOW (7.2%)[a]	LOW (5.2%)	LOW (9.0%)	MOD.LOW (28.6%)	LOW (7.1%)
2. Their Ability to Perform According to the Role	HIGH (4.54)[b]	HIGH (4.09)	MOD.HIGH (3.70)	MOD.HIGH (3.63)	MOD.HIGH (3.47)
3. Their Willingness to Perform According to the Role	HIGH (4.70)	HIGH (4.43)	HIGH (4.02)	HIGH (4.40)	HIGH (4.14)
4. Availability of Resources Necessary for Their Performance	MOD.HIGH (3.30)	MOD.LOW (2.70)	MOD.LOW (2.69)	MOD.LOW (2.52)	MOD.LOW (2.47)
5. Compatibility of the Existing Organization with the Innovation	HIGH (1.00)[c]	HIGH (1.02)	HIGH (1.00)	HIGH (1.00)	HIGH (1.03)

[a] Percentage of 56 items on which 80% or more of the paraprofessionals agreed were important or unimportant dimensions of their role. This factor included only paraprofessionals and the Ns were slightly different: Preprimary=13; Early Primary=60; Late Primary=11; Intermediate=12; and Junior High=26.

[b] Summary program scores for ability, willingness, and resources: LOW=1.00-1.99; MOD.LOW=2.00-2.99; MOD.HIGH=3.00-3.99; and HIGH=4.00-5.00.

[c] Summary program scores for compatibility: HIGH=1.00-1.25; MOD.HIGH=1.26-1.50; MOD.LOW=1.51-1.75; LOW=1.76-2.00.

TABLE 9

Ratings of Six Schools in the Early Primary Program on the Five Independent Variables of Paraprofessional Clarity, Willingness, and Ability, Availability of Necessary Resources, and Organizational Compatibility

Variable	School Ratings					
	A (N=5)	B (N=10)	C (N=9)	D (N=5)	E (N=11)	F (N=7)
1. Clarity of Paraprofessionals' Conception of Their Role.	LOW (10.0%)[a]	LOW (7.0%)	LOW (23.2%)	MOD.LOW (33.6%)	LOW (10.0%)	MOD.LOW (28.5%)
2. Their Ability to Perform According to the Role.	HIGH (4.47)[b]	MOD.LOW (2.91)	HIGH (4.61)	HIGH (4.38)	MOD.LOW (2.99)	HIGH (4.23)
3. Their Willingness to Perform According to the Role.	HIGH (4.65)	HICH (4.60)	HIGH (4.89)	HIGH (4.48)	HIGH (4.12)	HIGH (4.66)
4. Availability of Resources Necessary for Their Performance.	MOD.HIGH (3.40)	LOW (1.76)	MOD.HIGH (3.81)	LOW (1.85)	MOD.HIGH (3.15)	LOW (1.43)
5. Compatibility of the Existing Organization with the Innovation.	HIGH (1.13)[c]	HIGH (1.04)	HIGH (1.00)	HIGH (1.09)	HIGH (1.07)	HIGH (1.00)

[a] Percentage of 56 items on which 80% or more of the paraprofessionals agreed were important or unimportant dimensions of their role. Ns = 5 paraprofessionals in each of the six schools.

[b] Summary school scores for ability, willingness, and resources: LOW=1.00-1.99; MOD.LOW=2.00-2.99; MOD.HIGH=3.00-3.99; and HIGH=4.00-5.00.

[c] Summary school scores for compatibility: HIGH=1.00-1.25; MOD.HIGH=1.26-1.50; MOD.LOW=1.51-1.75; LOW=1.76-2.00.

TABLE 10

Ratings of Four Schools in the Junior High Paraprofessional Program on the Five Independent Variables of Paraprofessional Clarity, Willingness, and Ability of Staff, Availability of Necessary Resources, and Organizational Compatibility

Variable	School Ratings A (N=7)	B (N=11)	C (N=8)	D N=7)
1. Clarity of Paraprofessionals' Conception of Their Role.	LOW (7.1%)[a]	LOW (12.1%)	MOD.LOW (26.8%)	MOD.HIGH (51.8%)
2. Their Ability to Perform According to the Role.	MOD.HIGH (3.76)[b]	MOD.HIGH (3.55)	MOD.HIGH (3.98)	MOD.LOW (2.51)
3. Their Willingness to Perform According to the Role.	MOD.HIGH (3.96)	HIGH (4.22)	HIGH (4.54)	MOD.HIGH (3.73)
4. Availability of Resources Necessary for Their Performance.	MOD.LOW (2.64)	MOD.LOW (2.51)	MOD.LOW (2.72)	LOW (1.96)
5. Compatibility of the Existing Organization with the Innovation.	HIGH (1.01)[c]	HIGH (1.14)	HIGH (1.00)	HIGH (1.00)

[a] Percentage of 56 items on which 80% or more of the paraprofessionals agreed were important or unimportant dimensions of their role. Ns = 5 paraprofessionals in each of the four schools.

[b] School summary scores for ability, willingness, and resources: LOW=1.00-1.99; MOD.LOW=2.00-2.99; MOD.HIGH=3.00-3.99; and HIGH=4.00-5.00.

[c] School summary scores for compatibility: HIGH=1.00-1.25; MOD.HIGH=1.26-1.50; MOD.LOW= 1.51-1.75; and LOW=1.76-2.00.

TABLE 11

Predicted Rank Order Obtained from Ratings on the Five Independent Variables as Compared to Actual Rank Order According to Summary Implementation Scores

Program	Frequency of Ratings				Predicted Rank Order	Summary Scores and Assessed Rank Order[b]
	HIGH	MOD.HIGH	MOD.LOW	LOW		
Preprimary (N=11)	3	1	-	1	1	1 (4.04)[a]
Early Primary (N=70)	3	-	1	1	2	2 (3.97)
Late Primary (N=14)	2	1	1	1	4.5	4.5 (3.66)
Intermediate (N=16)	2	1	2	-	3	4.5 (3.65)
Junior High (N=33)	2	1	1	1	4.5	3 (3.81)

[a] Summary Scores on Implementation based on the averaged reports of teacher and paraprofessional observations. Lowest = 1.00 and the Highest = 5.00

[b] Spearman's rho = .775, p = .08 (One-tailed)

TABLE 12

Predicted Rank Order of 6 Schools with Early Primary Programs Obtained from Ratings on the Five Independent Variables as Compared to Actual Rank Order According to Summary Implementation Scores

School	Frequency of Ratings				Predicted Rank Order	Summary Scores and Assessed Rank Order[b]	
	HIGH	MOD.HIGH	MOD.LOW	LOW			
A (N=5)	3	1	-	1	1.5	1	(4.19)[a]
B (N=10)	2	-	1	2	6	5	(3.66)
C (N=9)	3	1	-	1	1.5	2.5	(4.00)
D (N=5)	3	-	1	1	3.5	4	(3.92)
E (N=11)	2	1	1	1	5	6	(3.60)
F (N=7)	3	-	1	1	3.5	2.5	(4.03)

[a] Summary Scores on Implementation based on the averaged reports of teacher and paraprofessional observations. Lowest = 1.00 and the Highest = 5.00

[b] Spearman rho = .872, $p < .02$ (one-tailed)

TABLE 12

Predicted Rank Order of 6 Schools with Early Primary Programs Obtained from Ratings on the Five Independent Variables as Compared to Actual Rank Order According to Summary Implementation Scores

School	Frequency of Ratings				Predicted Rank Order	Summary Scores and Assessed Rank Order[b]	
	HIGH	MOD.HIGH	MOD.LOW	LOW			
A (N=5)	3	1	-	1	1.5	1	(4.19)[a]
B (N=10)	2	-	1	2	6	5	(3.66)
C (N=9)	3	1	-	1	1.5	2.5	(4.00)
D (N=5)	3	-	1	1	3.5	4	(3.92)
E (N=11)	2	1	1	1	5	6	(3.60)
F (N=7)	3	-	1	1	3.5	2.5	(4.03)

[a] Summary Scores on Implementation based on the averaged reports of teacher and paraprofessional observations. Lowest = 1.00 and the Highest = 5.00

[b] Spearman rho = .872, $p < .02$ (one-tailed)

TABLE 13

Predicted Rank Order of 4 Schools with Junior High Paraprofessional Programs Obtained From Ratings on the Five Independent Variables as Compared to Actual Rank Order According to Summary Implementation Scores

School	Frequency of Ratings				Predicted Rank Order	Summary Scores and Assessed Rank Order[b]
	HIGH	MOD.HIGH	MOD.LOW	LOW		
A (N=7)	1	2	1	1	3.5	3.5 (3.54)[a]
B (N=11)	2	1	1	1	2	1.5 (4.15)
C (N=8)	2	1	2	-	1	1.5 (4.17)
D (N=7)	1	2	1	1	3.5	3.5 (3.52)

[a] Summary Scores on Implementation based on the averaged reports of teacher and paraprofessional observations. Lowest = 1.00 and the Highest = 5.00

[b] Spearman's rho = .995, p = .05 (one-tailed).